Access to in-situ Reference and Baseline Observations: status

C3S 311a Lot3

Fabio Madonna and the C3S 311a Lot3 team
fabio.madonna@imaa.cnr.it
Consiglio Nazionale delle Ricerche
Istituto di Metodologie per l'Analisi Ambientale
OBJECTIVES

Rationalize
Harmonize
Improve

Baseline and Reference in-situ Observations.

Quality control
Homogenization
Uncertainties

Surface Temperature
USCRN

T, RH, wind and IWV
GRUAN and IGRA
(radiosoundings), GNSS IGS

Trace Gases
NDACC, WOUDC, SHADOZ,
EuBrewNet, PGN, TCCON, ICOS


ORM11 Part-II, July 2021
SUB-CONTRACTORS

BIRA-IASB

Maynooth University
National University of Ireland Maynooth

National Physical Laboratory

UNIVERSITÀ DEGLI STUDI DI BERGAMO

WORLD METEOROLOGICAL ORGANIZATION

Universität Bremen

TALLINN UNIVERSITY OF TECHNOLOGY

ORM11 Part-II, July 2021
Scheme of the data fluxes at the level of the Data Management Facility (DMF), from the data suppliers to the C3S users.
INVENTORIES

At present time, inventories for the following networks have been provided:
GRUAN: 18 stations
IGRA: 2557 (including active and inactive stations).
NDACC O3: 197
NDACC CH4, CO: 53
TCCON CO, CO2, CH4: 32
PGN: 194
USCRN: 140
WOUDEC ozonesonde: 135
WOUDEC Dobson: 159
WOUDEC Brewer: 109
EuBrewNET: 46
SHADOZ: 17
GNSS: 657

Inventories include also a minimum amount of metadata (long, lat, station ID, WMO index, ECV, measurement type, instrument type, overlap among networks......).
DATASETS PROVIDED TO THE CDS: MAY 2021

1 dataset
- GRUAN: Submitted -> Reviewed -> Published

2 datasets
- IGRA
- RHAM: Submitted -> Under review by CDS

3 datasets
- WOUDC: Submitted -> Under review by CDS

1 dataset
- SHADOZ: Submitted -> Under review by CDS

4 datasets
- USCRN: Submitted -> Under review by CDS

9 datasets
- NDACC: Submitted -> Under review by CDS

ORM11 Part-II, July 2021
DATASETS PROVIDED TO THE CDS: MAY 2021

1 dataset
- TCCON: Software ready
  - Under implementation by C3S

2 dataset
- GNSS: Data ingestion

1 dataset
- EuBre: Submitted
  - Under review by CDS

3 dataset
- ICOS: Software tested and ready
  - Under implementation by C3S
Several tools for providing users with ancillary information or added-values products to fully exploit the data products have been designed and provided. For example, tool for temporal averaging and related uncertainty quantification.

Monthly averages of 30 years of Dobson data at Hohenpeissenberg
ANCILLARY PRODUCTS OZONE

Tools for temporal averaging and related uncertainty quantification

Monthly averages of 30 years of Dobson data at Hohenpeissenberg
Tools to quantify measurement representativeness

Example of ozone profile measurements made by a variety of techniques available in the NDACC data archive for Lauder, New Zealand (45°S, 170°W)

- ozonesondes
- lidar
- FTIR
- microwave radiometer
Data and metadata checks, quality checks (R and Pyhton scripts implemented by CNR)

Several quality checks are applied to each dataset before the delivery.
For example in the RHARM data processing, these are quickly summarized below:
- Metadata availability
- Compliancy with the CDM
- Physical plausibility check
- Completeness check (limitation to length of gaps in the time series)
- Threshold on the minimum amount of data available per month (15 values) to estimate uncertainties.
- Outliers identification/removal, using two criteria (e.g. mean-absolute-deviation, MAD, test and interquartile range test)
- Coherency checks.

These are in addition to the quality checks already applied by the networks.

Feedbacks are provided to the network Pis and a close interaction with each of them has been established to identify real issues and agree on a data handling strategy.
QUICK CHECKS: COVERAGE AND OUTLIERS

Potential outliers

Data coverage
COMMON DATA MODEL

In situ observation for all the C3S datasets are mapped to a Common Data Model (CDM).

«The CDM is a standardized and extensible collection of schemas (entities, attributes, relationships) that represents business concepts and activities with well-defined semantics, to facilitate data interoperability.»

It also defines the grammar to map several input datasets to unified definition of variable naming and units.

The CDM is implemented in recognition of the importance of a few existing standards:
• WIGOS MetaData Standard which extends the ISO19115 metadata standard, with additional mandatory elements describing both the station level and discovery metadata as well as specific information on the instrumentation used and processing steps
• BUFR
• CF conventions
Simplified schematic showing overview of C3S CDM-OBS (common data model for in situ obs.)
https://github.com/glamod/common_data_model/blob/master/cdm_full.pdf to be moved to C3S github page.
NEW COP2 PROPOSAL: C3S2 311 ITT

Title: Comprehensive Upper-Air, Baseline and Reference in situ observations

This tender widens the efforts spent within the C3S 311a Lot3 (“Access to observations from baseline and reference networks”) and C3S 311c Lot2 (“Historic In Situ Upper Air Database”) contracts of Cop1 to facilitate the access to and the homogenization of comprehensive upper-air, baseline and reference observations for a subset of GCOS relevant ECVs.

New datasets will be added to the C3S Climate Data Store, among the others PGN (Pandonia), NDACC NO2 and HCHO, BSRN.
GRUAN is the Upper-Air Reference network of GCOS.

- The GRUAN Lead Centre is already hosting a processing of the ozonesonde data.
- The goal is to provide a Reference product (i.e. traceable and quantified uncertainties).
- A proper treatment of uncertainties is still under discussion/implementation.
- GRUAN standards are expected to become consistent with the ASOPOS format.
- Eventual goal is for WOUDC / NDACC / SHADOZ / GRUAN ozonesondes to all provide identical profiles for the same flights, with interchangeable file formats (idealistic for now)
- File format converters already exist between these networks